

# Check dam extraction from remote sensing images using deep learning and geospatial analysis: A case study in the Yanhe River Basin of the Loess Plateau, China

SUN Liquan<sup>1,2,3</sup>, GUO Huili<sup>1,2</sup>, CHEN Ziyu<sup>4</sup>, YIN Ziming<sup>4</sup>, FENG Hao<sup>1,2</sup>, WU Shufang<sup>1,2,3\*</sup>, Kadambot H M SIDDIQUE<sup>5</sup>

<sup>1</sup> Key Laboratory of Agricultural Soil and Water Engineering in Arid and Semiarid Areas, Ministry of Education, Northwest A&F University, Yangling 712100, China;

<sup>2</sup> Institute of Water-Saving Agriculture in Arid Areas of China, Northwest A&F University, Yangling 712100, China;

<sup>3</sup> College of Water Resources and Architectural Engineering, Northwest A&F University, Yangling 712100, China;

<sup>4</sup> College of Natural Resources and Environment, Northwest A&F University, Yangling 712100, China;

<sup>5</sup> The UWA Institute of Agriculture and School of Agriculture & Environment, The University of Western Australia, Perth 6001, Australia

**Abstract:** Check dams are widely used on the Loess Plateau in China to control soil and water losses, develop agricultural land, and improve watershed ecology. Detailed information on the number and spatial distribution of check dams is critical for quantitatively evaluating hydrological and ecological effects and planning the construction of new dams. Thus, this study developed a check dam detection framework for broad areas from high-resolution remote sensing images using an ensemble approach of deep learning and geospatial analysis. First, we made a sample dataset of check dams using GaoFen-2 (GF-2) and Google Earth images. Next, we evaluated five popular deep-learning-based object detectors, including Faster R-CNN, You Only Look Once (version 3) (YOLOv3), Cascade R-CNN, YOLOX, and VarifocalNet (VFNet), to identify the best one for check dam detection. Finally, we analyzed the location characteristics of the check dams and used geographical constraints to optimize the detection results. Precision, recall, average precision at intersection over union (IoU) threshold of 0.50 (AP<sub>50</sub>), IoU threshold of 0.75 (AP<sub>75</sub>), and average value for 10 IoU thresholds ranging from 0.50–0.95 with a 0.05 step (AP<sub>50–95</sub>), and inference time were used to evaluate model performance. All the five deep learning networks could identify check dams quickly and accurately, with AP<sub>50–95</sub>, AP<sub>50</sub>, and AP<sub>75</sub> values higher than 60.0%, 90.0%, and 70.0%, respectively, except for YOLOv3. The VFNet had the best performance, followed by YOLOX. The proposed framework was tested in the Yanhe River Basin and yielded promising results, with a recall rate of 87.0% for 521 check dams. Furthermore, the geographic analysis deleted about 50% of the false detection boxes, increasing the identification accuracy of check dams from 78.6% to 87.6%. Simultaneously, this framework recognized 568 recently constructed check dams and small check dams not recorded in the known check dam survey datasets. The extraction results will support efficient watershed management and guide future studies on soil erosion in the Loess Plateau.

**Keywords:** check dam; deep learning; geospatial analysis; remote sensing; Faster R-CNN; Loess Plateau

**Citation:** SUN Liquan, GUO Huili, CHEN Ziyu, YIN Ziming, FENG Hao, WU Shufang, Kadambot H M SIDDIQUE. 2023. Check dam extraction from remote sensing images using deep learning and geospatial analysis: A case study in the Yanhe River Basin of the Loess Plateau, China. Journal of Arid Land, 15(1): 34–51. <https://doi.org/10.1007/s40333-023-0091-7>

\*Corresponding author: WU Shufang (E-mail: wsfjs@163.com)

Received 2022-09-01; revised 2022-12-06; accepted 2022-12-28

© Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Science Press and Springer-Verlag GmbH Germany, part of Springer Nature 2023

## 1 Introduction

Check dams, one of the most effective soil and water conservation engineering measures for trapping sediments and mitigating soil erosion effects, are used worldwide (Abbasi et al., 2019; Rahmati et al., 2020; Lucas-Borja et al., 2021). They are constructed in gullied channels to mitigate flood damage (Yazdi et al., 2018), control sediment transport (Shi et al., 2019), stabilize slopes and torrential channels (Piton and Recking, 2017), and serve as high-quality cropland when filled with sediment (Jin et al., 2012). More than 110,000 check dams have been built on the Loess Plateau in the last 50 years (Wang et al., 2011), with most now abandoned or no longer maintained. In 2011, 58,446 check dams remained on the Loess Plateau, with 927.6 km<sup>2</sup> available for cropland (Liu, 2013).

Despite major efforts to implement check dam construction to control soil erosion on the Loess Plateau (Wang et al., 2021), problems remain. Some of the early built check dams in remote regions with few inhabitants have lost their function (Jin et al., 2012) due to the lack of reasonable management and maintenance, and breakage during rainstorms could cause more damage than normal soil losses (Bai et al., 2020). Therefore, obtaining accurate information on the number of existing check dams, their location, and spatial distribution is vital for analyzing their effects on erosion reduction, timely maintaining and consolidating, and planning suitable dam sites in future (Shi et al., 2019; Pourghasemi et al., 2020). Traditionally, the distribution of check dams comes from documented construction data and in situ hydrographic surveys. However, these methods are usually time-consuming, labor-intensive, and costly. With the development of remote sensing technology, object extraction from remote sensing images is possible, providing invaluable and timely information on spatial and spectral attributes of check dams to support detection and monitor tasks (Tian et al., 2013; Mi et al., 2015). Zhao (2007) used a pixel-based method (supervised classification) to extract dam areas from high-resolution remote sensing images. Hou (2013) used object-based image analysis (OBIA) to automatically extract check dams by considering the texture characteristics of dam land and water body parts. Alfonso-Torreno et al. (2019) identified check dams with high-resolution aerial photographs captured from an Unmanned Aerial Vehicle (UAV) and estimated the volume of sediments deposited in those check dams. These studies focused on extracting dam cropland or water bodies controlled by check dams in a small watershed; however, research on check dam identification and distribution using remote sensing images in broad regions is rare. Moreover, as remote sensing images are complex, traditional image processing methods have become less effective or failed in robustly processing large datasets (Kamilaris and Prenafeta-Boldu, 2018).

In recent years, the rapid development of artificial intelligence, especially deep learning methods in the computer vision field, has brought new opportunities for high-resolution remote sensing image analysis. Compared with traditional machine learning methods, deep learning based on convolutional neural networks (CNNs) has strong feature extraction capability and high accuracy (Ghanbari et al., 2021), with great potential for application in areas such as regional-scale land use classification and ground object identification and extraction (Mahdianpari et al., 2018; Khelifi and Mignotte, 2020; Konstantinidis et al., 2020). However, our literature review found few studies focusing on check dam detection using CNNs. Li et al. (2021) proposed a check dam extraction method that integrates OBIA and a U-Net deep learning semantic segmentation model to detect areas for check dams but not specific check dam locations. Object detection based on CNNs, although not applied to check dam identification in broad areas, has been used successfully for many other target recognition tasks in the remote sensing field (Fu et al., 2019), including building, ship, airplane, and airport detection, and precision agriculture (Apolo-Apolo et al., 2020; Reda and Kedzierski, 2020; Mur et al., 2022). Ding et al. (2018) improved the Faster R-CNN with enhanced Visual Geometry Group (VGG) 16-Net and tested it using remote sensing datasets of aircraft and automobiles; the results showed that the proposed approach could accurately and efficiently detect objects. Wu et al. (2021) combined the local fully convolution neural network (FCN) and You Only Look Once (version 5) (YOLOv5) to detect

small targets in remote sensing images, reporting more accurate feature recognition and detection performance for densely arranged target images.

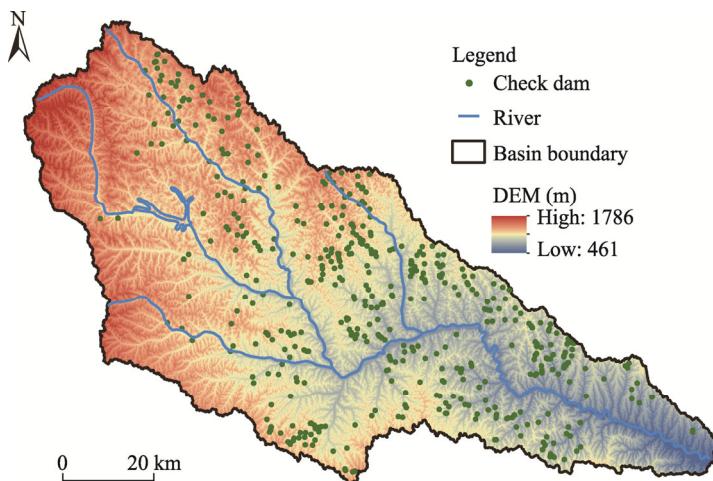
There will inevitably be misidentified objects when using target detection models, considering the complexity of remote sensing images. Therefore, post-processing detection also needs to be improved. Li et al. (2021) proposed a workflow for detecting unknown airport distributions in a broad region based on deep learning and geographic analysis, performing a spatial analysis using geographical data such as road networks and water systems to achieve fast and reliable airport detection. Spatial analyses could help us analyze the characteristics of ground objects and their relationships and solve complex location-oriented problems, which lends new perspectives to decision-making. Geographical factors, such as water bodies, land cover, slope, topography, and gully width, significantly impact the construction of check dams. Therefore, we hypothesized that introducing a geospatial analysis approach to identify check dams would greatly improve the reliability of the results.

In this research, we aimed to develop a check dam detection framework using deep learning and geospatial analysis to identify check dams from high-resolution remote sensing images at a regional scale. The specific objectives were to (1) compare the performance of different object detectors based on deep learning and determine the optimal detector for check dam identification; and (2) optimize the detection results from deep learning by conducting geospatial analysis and comprehensive discrimination based on open-source remote sensing products. The above research would provide data support for researchers to assess hydrological and ecological effects quantitatively and for watershed managers to plan the layout of check dams in future. Meanwhile, the proposed method offers fast, automatic, and low-cost detection for supervising check dams on the Loess Plateau, especially in broad areas where economic conditions impede ground monitoring.

## 2 Materials and methods

### 2.1 Study area

This study uses the Yanhe River Basin (Fig. 1), located in the hilly and gully region of the Loess Plateau, China ( $36^{\circ}22'–37^{\circ}20'N$ ,  $108^{\circ}39'–110^{\circ}29'E$ ), as a case study. The Yanhe River is a first-order tributary (about 286.9 km long) of the Yellow River, covering a drainage area of 7725  $km^2$ . The altitude of the basin ranges from 497 to 1777 m, decreasing gradually from the northwest to the southeast. The Yanhe River Basin contains thick loess, a fine silt soil that is loose and weakly resistant to raindrop erosion and runoff scouring. The climate is a continental semiarid monsoon with average annual precipitation of 500–550 mm. However, the precipitation



**Fig. 1** Overview of the Yanhe River Basin and spatial distribution of partial check dams in the study area. DEM, Digital Elevation Model.

varies seasonally and is extremely uneven, with more than 70% of the annual precipitation occurring as short-duration, high-intensity rainstorms in summer from June to September (Bai et al., 2019), causing severe soil erosion and degrading the landform.

Since the 1950s, various soil and water conservation measures have been conducted in the Yanhe River Basin, mainly check dams and afforestation. The construction of check dams has dramatically reduced soil erosion rates and trapped thousands of tons of sediment, significantly decreasing the sediment load at the Ganguyi station (Wei et al., 2018). The Yellow River Conservancy Committee reported approximately 800 large and medium check dams by the end of 2008 (Sun and Wu, 2022), and the number has continued to increase.

## 2.2 Remote sensing data preprocessing

The data used in this study include GaoFen-2 (GF-2) and Google Earth remote sensing images, ASTER Global Digital Elevation Model (DEM) v3, and the European Space Agency (ESA) WorldCover 10 m 2020 v100 (Zanaga et al., 2021), as shown in Table 1.

**Table 1** Basic information on data used in this study

Data type	Product name	Data time	Spatial resolution (m)	Source
Remote sensing image	GaoFen-2	2020	PMS: 4.0; PAN: 1.0	Land Observation Satellite Data Service Platform of the China Resources Satellite Application Center
	Google Earth image	2020	0.3 and 1.0	Google Earth
DEM	ASTER Global DEM v3	2019	30.0	National Aeronautics and Space Administration Earthdata
Land cover	ESA WorldCover 10 m 2020 v100	2020	10.0	ESA

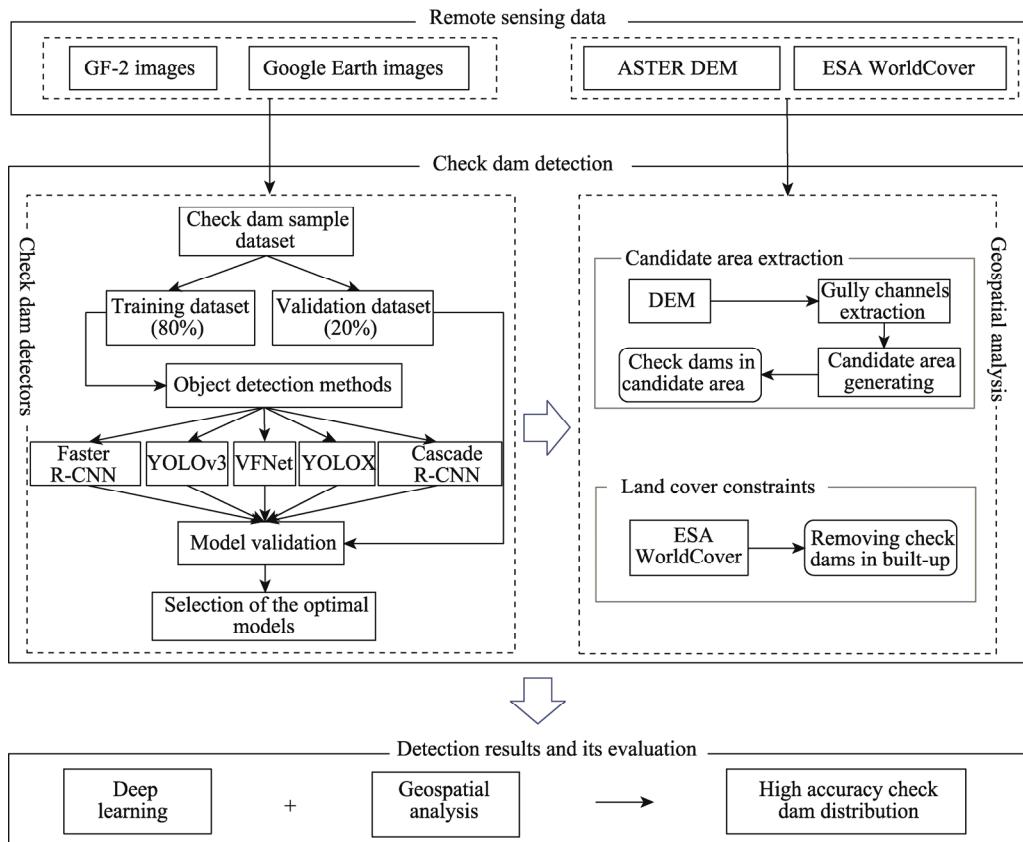
Note: DEM, Digital Elevation Model; ESA, the European Space Agency; PMS, multispectral band; PAN, panchromatic band.

A total of 20 GF-2 images covering the Yanhe River Basin were collected as the main data source, which have a resolution of 1.0 m in the panchromatic band and 4.0 m in multispectral band (blue, green, red, and near-infrared spectrum) on a swath of 45 km. To avoid the effect of snow and clouds on identifying check dams, we acquired these images in different seasons (April to November), including three images on 5 April 2020, four images on 25 April 2020, two images on 24 May 2020, six images on 15 September 2020, and five images on 19 October 2020, with less than 5% cloud coverage in each scene image. All GF-2 images were preprocessed using Environment for Visualizing Images (ENVI) software (version 5.3.1). We used DEM to perform rational polynomial coefficients (RPC) orthorectification on the multispectral and panchromatic images, projecting them into Universal Transvers Mercator coordinate system. Next, the multispectral image was registered to the corresponding panchromatic image using polynomial warping with automatically generated tie points, providing a registration error of less than 1 pixel. Subsequently, the Gram-Schmidt Pan Sharpening method was applied to fuse panchromatic and multispectral bands, enhancing the spatial resolution of multispectral bands from 4.0 to 1.0 m (Laben and Brower, 2000). Finally, the bit depth of all fused images was unified to 8 bits using optimized linear stretch. In addition, some Google Earth images with a spatial resolution of 0.3 and 1.0 m were acquired as supplementary data for areas not covered by GF-2 images. Rich image variations in different seasons and sources can also overcome the shortcomings of insufficient image diversity and target variability, improving the robustness and generalization ability of the model.

## 2.3 Methodology

Figure 2 illustrates the framework of the proposed check dam detection method: (1) remote sensing dataset preparation, (2) check dam detection based on deep learning object detection models, and (3) geospatial analysis and comprehensive discrimination for results acquired from step (2). While deep learning object detection models can identify targets quickly and accurately,

there are inevitably errors when identifying features from complicated remote sensing images in broad regions due to computing hardware limitations. To solve this, we used the sliding window (1024×1024 pixel) method when detecting check dams in the Yanhe River Basin. A non-maximum suppression algorithm was used to filter the redundant detection boxes and optimize the detection results.



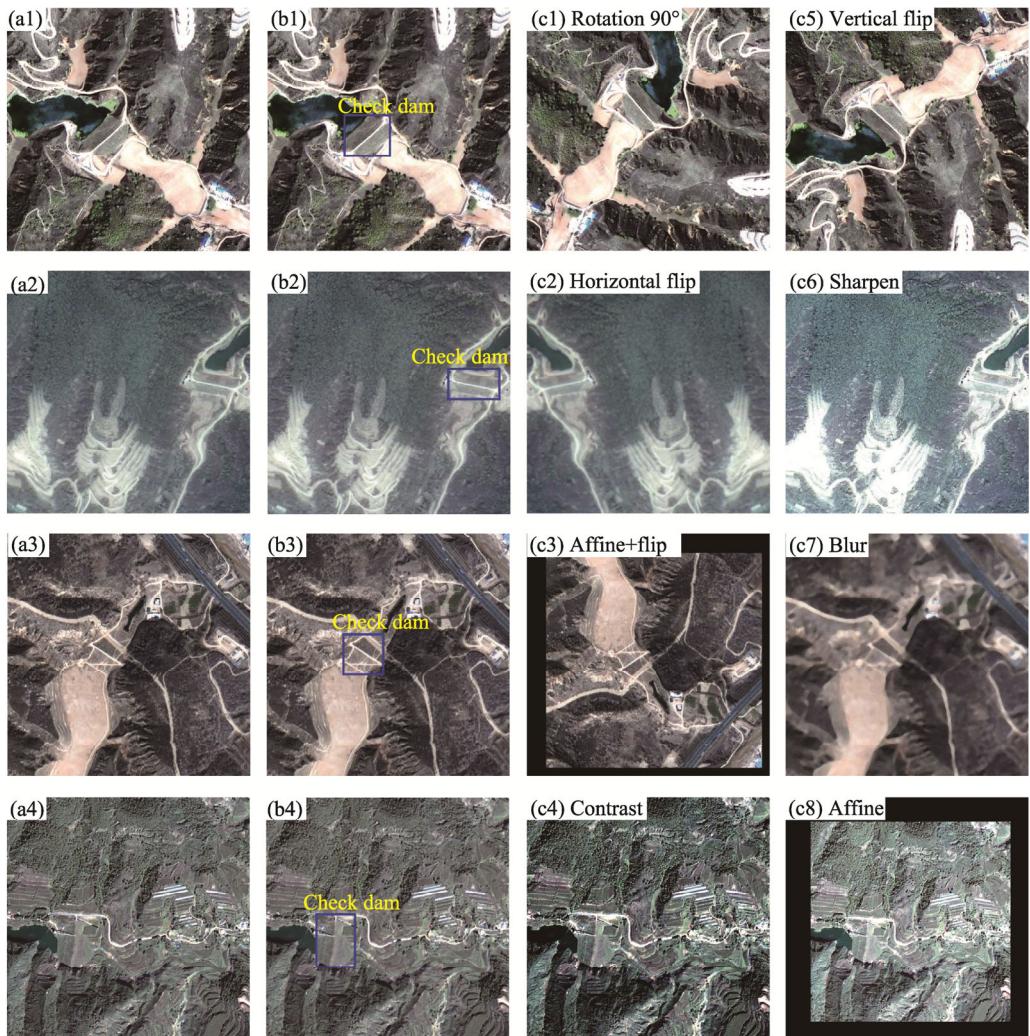
**Fig. 2** Technical workflow for check dam detection across broad regions. GF-2, Gaofen-2; ASTER DEM, ASTER Global Digital Elevation Model (DEM) v3; ESA WorldCover, the European Space Agency WorldCover 10 m 2020 v100; YOLOv3, You Only Look Once (version 3); VFNet, VarifocalNet.

### 2.3.1 Dataset preparation

We marked more than 600 large and medium check dams in the Yanhe River Basin and surrounding areas using survey data from the Bulletin of First National Census for Water in China (Ministry of Water Resources of China, 2013) and field data of check dams in Baota District and Yanchang County of Shaanxi Province conducted by the Water Conservancy Bureau in 2015 (Fig. 1). We conducted field surveys of check dams in October 2021 with Unmanned Aerial Vehicles (UAVs) and handheld Global Positioning System receivers (Trimble Juno 3D, Shenzhen Pengjin Technology Co. Ltd., Shenzhen, China) to confirm the reliability of collected check dam data.

After acquiring the required remote sensing images, including GF-2 and Google Earth images, we prepared image datasets for training. The morphological characteristics of check dams are relatively simple in remote sensing images, as they are easily recognized, especially dam bodies and dam land. The dam slope usually presents a rectangle or quasi-rectangle in the image, with quasi-triangles in a few cases. The dam crest often plays the role of road and bridge to connect the traffic on both sides of the gully in a linear feature. Dam land is often formed by intercepted sediment and water bodies; when the dam fills with sediment, the resulting flat land becomes cropland for agriculture, such that the dam land is flat compared with the surrounding terrain in

the image. The images with identified check dams were subset to  $1024 \times 1024$  pixel sub-images with 25% overlap to speed up the training process and improve hardware usage efficiency before annotating with ArcGIS Pro software. All images were confirmed for the presence of check dams using the survey data list, with check dams marked with rectangular boxes (Fig. 3). The overlap avoids detecting borders when check dams are only partially contained in the sub-image. A total of 1326 images containing check dam annotations were acquired. In practice, training a good deep learning model requires many samples. We enhanced the size of the dataset using data augmentation techniques to reduce network overfitting and obtain a strongly generalizable model (Fig. 3). New images and annotations were generated using a random combination of rotating, flipping, adding noise, blurring and resizing the original images, and changing colors using a Python script. The samples were enhanced about 10 times. We subsequently constructed a dataset of 12,988 samples, divided into an 8:2 ratio comprising 10,392 training samples and 2596 validation samples.



**Fig. 3** Display of the original images (a), images with annotations (b), and augmented images (c). (a1–a4), original images of check dams; (b1–b4), images of check dams with annotation; (c1–c8), images of check dams after data augmentation.

### 2.3.2 Deep learning network for check dam detection

We used MMDetection, an object detection toolbox containing a rich set of object detection and

instance segmentation networks, to rapidly build the desired deep learning object detection models on the PyTorch open-source deep learning framework (Chen et al., 2019).

Generally, existing deep learning methods designed for object detection can be divided into region proposal-based methods and regression-based methods. Region proposal-based detectors, such as R-CNN, Fast R-CNN, and Faster R-CNN, explicitly extract bounding box candidates and separately classify candidate-related features (Ren et al., 2017). Regression-based detectors, such as You Only Look Once (YOLO), Single Shot MultiBox Detector (SSD), and RetinaNet, unify candidate region detection and feature classification (Fu et al., 2019). Here, we selected the most representative region proposal-based detectors, including Faster R-CNN and Cascade R-CNN, and regression-based detectors, including You Only Look Once (version 3) (YOLOv3), YOLOX, and an intersection over union (IoU)-aware dense object detector (VarifocalNet; abbreviated as VFNet), to assess their ability to detect check dams (Cai and Vasconcelos, 2017; Ren et al., 2017; Redmon and Farhadi, 2018; Ge et al., 2021; Zhang et al., 2021). These networks have been successful for other target recognition tasks in the remote sensing field (Fu et al., 2019), but they are rarely applied to check dam detection in broad areas. Moreover, Feature Pyramid Networks (FPN) were added to these networks to solve the multi-scale problem in check dam detection and improve the performance of check dam detection.

Faster R-CNN is a two-stage target detection network. In the first stage of check dam identification, the detector extracts feature maps by convolutional neural network (CNN) backbone from remote sensing images before inputting the feature maps into the region proposal network (RPN) to generate region proposals. The second stage calculates classification and coordinate regressions to region proposals to predict the border of the check dam location and its confidence level, requiring an IoU threshold to define positives and negatives. A detector trained with low IoU threshold (e.g., 0.5 in the Faster R-CNN) usually produces noisy detections. However, detection performance tends to degrade with increasing IoU thresholds. The Cascade R-CNN, an improvement network based on Faster R-CNN, is proposed to address these problems. It comprises a sequence of detectors trained with increasing IoU thresholds (0.5, 0.6, and 0.7 in this study) to be sequentially more selective against close false positives. ResNeXt-101 is selected as the backbone for the feature extraction of check dams in complicated remote sensing images.

YOLOv3 is a representative regression-based object detection method. It works by resizing the input images to  $608 \times 608$  pixel and using the Darknet53 backbone to perform feature extraction. This backbone sets up links between layers and skips some convolution layers to avoid the vanishing gradient problem. The images were down-sampled 32 times, with scaled feature maps ( $19 \times 19$ ,  $38 \times 38$ , and  $76 \times 76$ ) obtained and used to detect small, medium, and large targets. Meanwhile, the deeper feature maps were up-sampled twice and merged with the shallower feature maps. We divided the input images into default grids according to the scale of feature maps. Anchor boxes obtained by K-means clustering were tiled onto each grid cell, and predictions of bounding boxes, confidences, and object names were made accordingly. Ge et al. (2021) used YOLOv3 as a baseline and proposed YOLOX, which integrates excellent advantages, including decoupled head, mosaic data enhancement, SimOTA, and anchor-free mechanism, to improve model performance. Compared with YOLOv3, You Only Look Once (version 4) (YOLOv4), and YOLOv5, YOLOX not only has a simpler structure but also exhibits good inference speed and detection accuracy, which is advantageous in the context of small target detection. The YOLOX comprises three main parts, i.e., backbone, neck, and YOLO head. Three feature layers are extracted in the CSPDarknet backbone and then fused in the neck part. The YOLO head includes a classifier and a regressor to judge feature points and determine whether objects correspond to them.

VFNet is a new object detection method for accurately ranking a huge number of candidate detections based on IoU-aware Classification Score (IACS), which can simultaneously represent the confidence of object presence and localization accuracy for grading the detection. VFNet contains a new loss function, Varifocal Loss, for training a dense object detector to predict IACS, a new efficient star-shaped bounding box feature representation for estimating IACS and refining

coarse bounding boxes, and fully convolutional one-stage object detection+adaptive training sample selection (FCOS+ATSS) architecture. It uses the varifocal loss to predict IACS for each image. We used Res2Net-101 as the backbone to extract features of the input images, and then the feature pyramid network to generate five feature maps at different scales. Lastly, we performed bounding box regression prediction and fine-tuning refinement in the VFNetHead network.

The detection models were trained and validated on a workstation with an Intel Core i7-7700 central processing unit (CPU) and NVIDIA RTX Tesla P100 (16 GB) general processing unit (GPU) running on an Ubuntu 18.04 system. Table 2 shows the hyper-parameters applied in the experimental configurations for the training object detectors to achieve the highest model performance in terms of accuracy.

Training a CNN from scratch is computationally expensive and time-consuming. Therefore, transfer learning was used to transfer the knowledge learned from one model trained on a large dataset, such as Microsoft COCO (MSCOCO), to another model to solve a specific task (Chen et al., 2018). We used transfer learning to train the check dam detection models based on pre-trained backbone networks in the MSCOCO dataset.

**Table 2** Hyper-parameters used for training deep learning object detection networks

Model	Faster R-CNN	Cascade R-CNN	YOLOv3	YOLOX	VFNet
Backbone	ResNeXt-101	ResNeXt-101	Darknet-53	CSPDarknet	Res2Net-101
Batch	2	2	12	4	2
Optimizer	SGD	SGD	SGD	SGD	SGD
Momentum	0.90	0.90	0.90	0.90	0.90
Weight decay	0.0001	0.0001	0.0005	0.0005	0.0001
Base learning rate	0.00250	0.00250	0.00100	0.00125	0.00125
Epoch	24	24	30	30	24

Note: YOLOv3, You Only Look Once (version 3); VFNet, VarifocalNet; SGD, stochastic gradient descent.

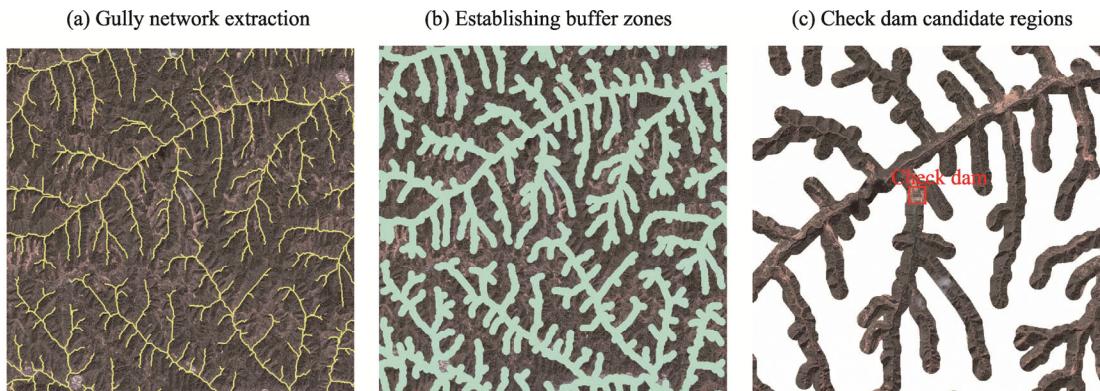
### 2.3.3 Geospatial analysis

We introduced geospatial analysis methods to improve the precision of check dam identification. Based on the topographic conditions of check dam construction and land cover types in the check dam areas, we extracted the corresponding candidate regions in gullies and land cover types.

According to the location feature that the check dam is constructed along gullies and channels, we extracted the candidate areas of check dam identification from DEM in ArcGIS 10.2 (Fig. 4) to filter incorrect detection results obtained from deep learning models and improve the accuracy of check dam identification. There are two major steps for candidate area extraction: (1) extract the gully network using the D8 algorithm (Ngula Niipele and Chen, 2019) in ArcHydro tools of ArcGIS 10.2; the workflow includes DEM reconditioning, depression filling, flow direction, flow accumulation calculations, and gully network generating. This stage identifies the optimal gully network as long as the drainage lines containing fewer branches can pass through all check dams in the study area. We used different thresholds, including 50, 100, 200, 300, 500, and 1000 for flow accumulation cut-off values to extract gully networks (Fig. S1), determining 200 as the most suitable threshold; and (2) establish buffer zones of the gully network. Based on the field survey, we determined the specified distance (135 m) for creating buffer zones around the gully network that covers the check dams in the study area by combining their distribution and scale.

In the early stage after construction, check dams are used mainly to retain rainstorm-caused runoff, intercept sediment, and generally form a water body behind the dam. When the check dams fill with sediment, the generated flat land can be used for agricultural production due to its humus-rich soil carried by runoff. There are eight land cover classes in the Yanhe River Basin

according to ESA WorldCover 10 m 2020 v100: tree cover, shrubland, grassland, cropland, built-up, bare or sparse vegetation, permanent water bodies, and herbaceous wetland. The land cover type analysis of the checks dams collated in the survey data revealed six main land cover classes: cropland, water bodies, bare land, shrubland, tree cover, and grassland. Thus, we used land cover constraints to refine the detection results by deleting detection boxes not located in such land cover types.



**Fig. 4** Check dam candidate area extraction in ArcGIS 10.2. (a), gully network extraction from ASTER Global Digital Elevation Model (DEM) v3 using the D8 algorithm; (b), establishing buffer zones around the gully network with the specified distance of 135 m; (c), check dam candidate regions extraction based on buffer zones.

### 2.3.4 Model performance evaluation

The performance of each model was evaluated using precision, recall, precision-recall (P-R) curve, average precision (AP), and inference speed (FPS). Precision refers to the ratio of the number of correct detections to the total number of detections. Recall refers to the ratio of the number of correct check dam detections to the total ground truth in the validated dataset. The P-R curve shows the precision and recall at different IoU thresholds. When evaluating models, if the IoU between the ground truth and the detecting bounding box exceeded a predefined threshold  $\lambda$  (Eq. 1), the detection was noted as a true positive; otherwise, the detection was noted as a false positive. The AP is the area under the P-R curve, a standard for evaluating the precision of the deep learning object detection models (a higher AP value represents higher detection accuracy). The AP calculation in this study was based on the evaluation criteria of the MSCOCO dataset (Lin et al., 2014), including IoU threshold of 0.50 ( $AP_{50}$ ), IoU threshold of 0.75 ( $AP_{75}$ ), and average value for 10 IoU thresholds ranging from 0.50–0.95 with a 0.05 step ( $AP_{50-95}$ ). The FPS is the number of images detected per second or the time to detect each image. Precision, recall, and AP are calculated as follows:

$$IoU = \frac{\text{area}(\text{detection} \cap \text{ground truth})}{\text{area}(\text{detection} \cup \text{ground truth})} > \lambda, \quad (1)$$

$$P = \frac{TP}{TP + FP} = \frac{TP}{N}, \quad (2)$$

$$R = \frac{TP}{TP + FN}, \quad (3)$$

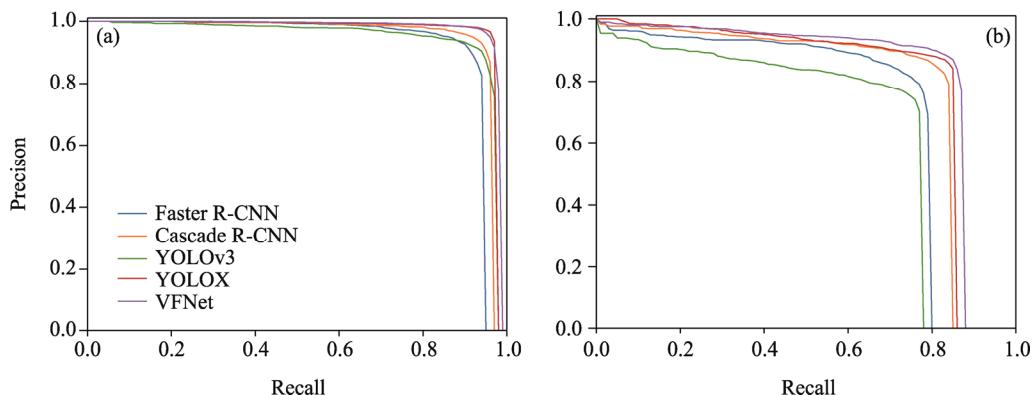
$$AP = \int_0^1 P(R) dR, \quad (4)$$

where  $IoU$  is the intersection over union;  $\lambda$  is the predefined threshold;  $P$  is the precision;  $TP$  is the number of check dams correctly detected by the models;  $FP$  is the number of false detections;  $N$  is the number of all detected check dams;  $R$  is the recall; and  $FN$  is the number of missed detections.

### 3 Results

#### 3.1 Comparing the performance of five models for check dam identification

Figure 5 shows the P-R curves of the five methods at different IoU thresholds; the area under the curve is the AP value of the corresponding model (Table 3). The P-R curve of VFNet completely enclosed those of the others regardless of whether the IoU threshold was 0.50 or 0.75, suggesting that VFNet has optimal performance for check dam identification, followed by YOLOX and Cascade R-CNN. YOLOv3 significantly outperformed Faster R-CNN when the IoU threshold was 0.50, but its performance decreased significantly as the IoU threshold increased; for example, at an IoU threshold of 0.75, YOLOv3 performed the worst in terms of recognition ability.



**Fig. 5** Comparison of precision-recall (P-R) curves for the five deep learning models at different intersection over union (IoU) thresholds. (a), P-R curves at  $\text{IoU}=0.50$ ; (b), P-R curves at  $\text{IoU}=0.75$ .

Table 3 shows that all models except YOLOv3 reached 60.0%, 90.0%, and 70.0% of the AP values at IoU thresholds of 0.50:0.95, 0.50, and 0.75, respectively. Among the object detection deep learning networks, VFNet had the highest AP (69.9%) in the validation datasets. Cascade R-CNN and YOLOX improved the AP value by 5.5% and 11.3%, respectively, compared to Faster R-CNN and YOLOv3, indicating that the improvements in Faster R-CNN and YOLOv3 greatly enhanced model performance, especially for YOLOX. In addition, we compared the inference speed of each model (image size:  $1024 \times 1024$ ). YOLOv3 was the fastest, with an inference speed greater than 25.0 image/s. Even though the number of weight parameters increased compared to YOLOv3, the detection speed of YOLOX still reached 15.8 image/s. The inference speed of Faster R-CNN and Cascade R-CNN was only about 4.0 image/s, while the speed of VFNet slightly improved, reaching 5.5 image/s.

**Table 3** Comparison of average precision and inference speed for different object models

Model	Backbone	Average precision (%)			Inference speed (image/s)
		$\text{IoU}=0.50:0.95$	$\text{IoU}=0.50$	$\text{IoU}=0.75$	
Faster R-CNN	ResNeXt-101	61.5	92.5	72.3	4.2
Cascade R-CNN	ResNeXt-101	67.0	94.9	78.5	3.9
YOLOv3	Darknet-53	58.1	94.3	66.5	25.4
YOLOX	CSPDarknet	69.4	96.4	80.3	15.8
VFNet	Res2Net-101	69.9	97.2	82.5	5.5

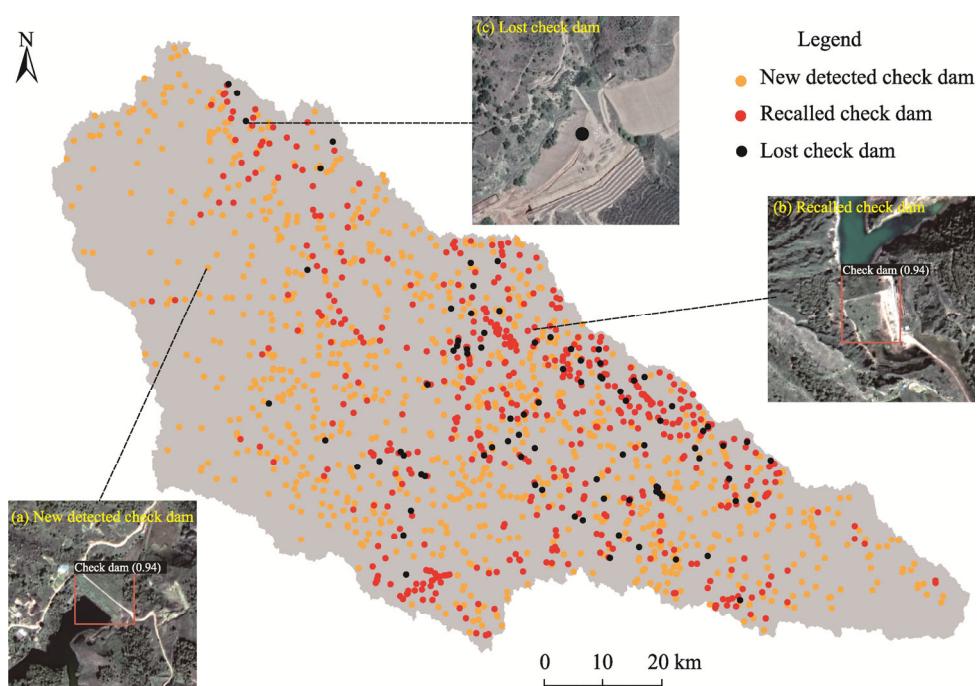
Note: IoU represents intersection over union.

Overall, the five models achieved good results for check dam detection. Considering the detection precision and speed (Table 3), YOLOX outperformed the other models with superior

performance for both accuracy and efficiency. However, detection accuracy is more important than speed for check dam identification, so VFNet was also selected to detect check dams in this study.

### 3.2 Check dams detection based on optimal detectors and geographical analysis

After identifying the optimum object detection models, we integrated VFNet and YOLOX to perform check dam identification on the 20 GF-2 remote sensing images covered on the Yanhe River Basin, retaining the detection boxes with confidence thresholds more than 0.5. This process identified 1390 detection boxes (Fig. 6). We validated the detection results using check dam survey data, field investigation, and visual judgment of the available high-resolution historical images on Google Earth by the experts in the field of earth observation interpretation according to the interpretation symbol of check dams depicted in Section 2.3. As a result, 1092 detection boxes were identified correctly as check dams, and 298 detection boxes were misidentified (precision: up to 78.6%). According to the check dam survey data mentioned above, we identified 602 check dams distributed in the Yanhe River Basin, of which 524 were recalled (recall rate: up to 87.0%), and 78 were not recognized (Table 4).



**Fig. 6** Results of check dam detection based on VFNet and YOLOX in the Yanhe River Basin. (a), the image of new detected check dam; (b), the image of recalled check dam; (c), the image of lost check dam. Detection results in (a) and (b) show the predicted bounding box of check dam and its corresponding confidence score.

**Table 4** Evaluation of check dam detections after geospatial analysis and comprehensive discrimination in the Yanhe River Basin

Method	Precision evaluation			Recall evaluation		
	Detection box	Correct detection	Precision (%)	Validation data of check dams	Recalled check dams	Recall (%)
DL	1390	1092	78.6	602	524	87.0
DL+GA	1243	1089	87.6		521	86.5

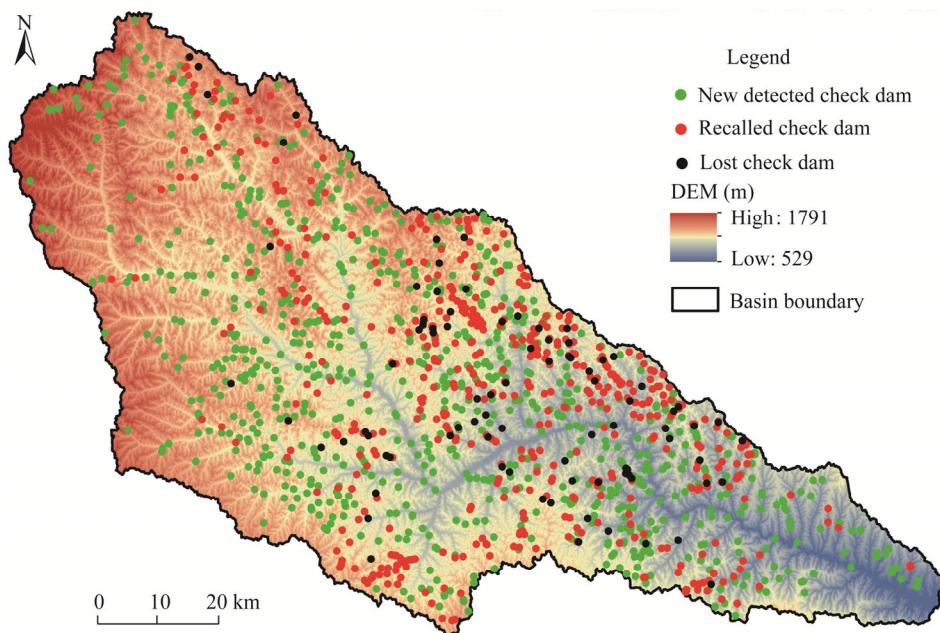
Note: DL represents deep learning object detect methods, including VFNet and YOLOX; GA represents geospatial analysis.

Table 4 shows the results of the geospatial analysis. We used the candidate regions acquired from DEM (Fig. 4) and land cover from ESA WorldCover 10 m 2020 v100 as restrictive

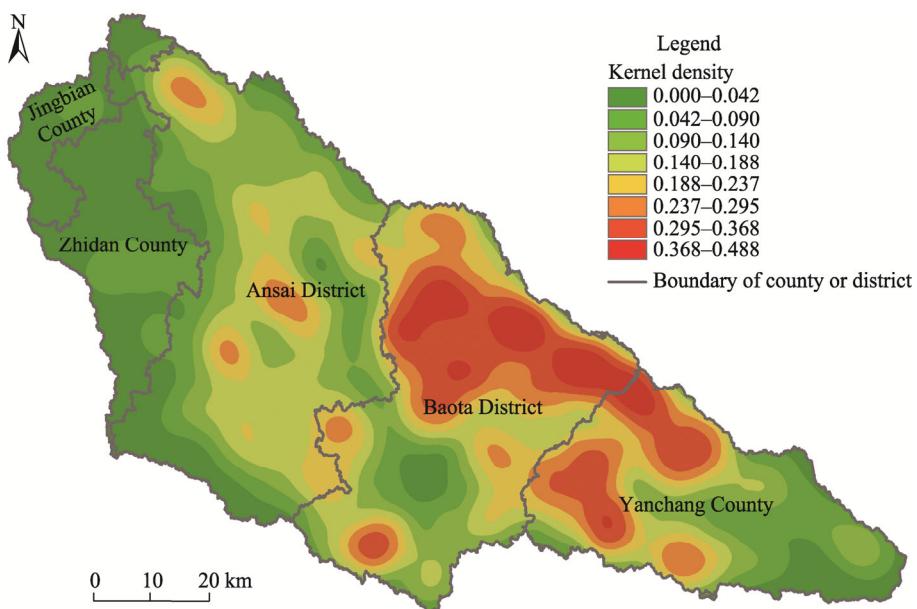
conditions to reduce incorrect detections. Finally, we removed 147 incorrect detections from the low-confidence check dam detection results recognized by deep learning and obtained 1243 detections as high-confidence check dam detection results, with 1089 correctly identified as check dams. The geospatial analysis and comprehensive discrimination removed about 50.0% of incorrect detections. The precision of check dam identification improved by 9.0%, reaching 87.6%, and the recall rate reached 86.5%. Simultaneously, 568 check dams, including recently constructed and those not recorded in the known check dam survey datasets, were recognized by our proposed framework. However, due to the limited accuracy of land cover, incorrectly applying the spatial analysis eliminated three detections that the deep learning models correctly recognized. Thus, the framework could rapidly and precisely detect check dams and provide location and distribution information of recently constructed check dams to complement the survey data.

### 3.3 Check dams distribution in the Yanhe River Basin

Based on the proposed framework, we extracted check dams in the Yanhe River Basin (Fig. 7). We used the 'Kernel Density' tool from ArcGIS 10.2 to generate a density map of check dams in the Yanhe River Basin (Fig. 8) to show the spatial distribution of constructed check dams, and provide reference for macroscopically planning the layout of check dams in future for watershed managers. The density map was classified into several categories using the natural breakpoint method, with the boundary divided at the position with large numerical differences. As a result, there are noticeable regional differences in the spatial distribution of check dams within the Yanhe River Basin. Check dams are mainly concentrated in the central and northeastern parts of the study region, with higher density values ranging from 0.300 to 0.500 (Fig. 8). Two agglomeration areas of high density are mainly located in the Baota District and the western part of Yanchang County near the Baota District, which may be because the Baota District acts as the administrative center of Yan'an City of Shaanxi Province, with an important role in culture and economic development. Therefore, plenty of check dams have been constructed in this region to regulate runoff and control soil erosion. However, the Ansai District has medium-density values, and the remaining areas have a low degree of agglomeration of check dams.



**Fig. 7** Check dam detection results based on the integration method of deep learning and geospatial analysis in the Yanhe River Basin



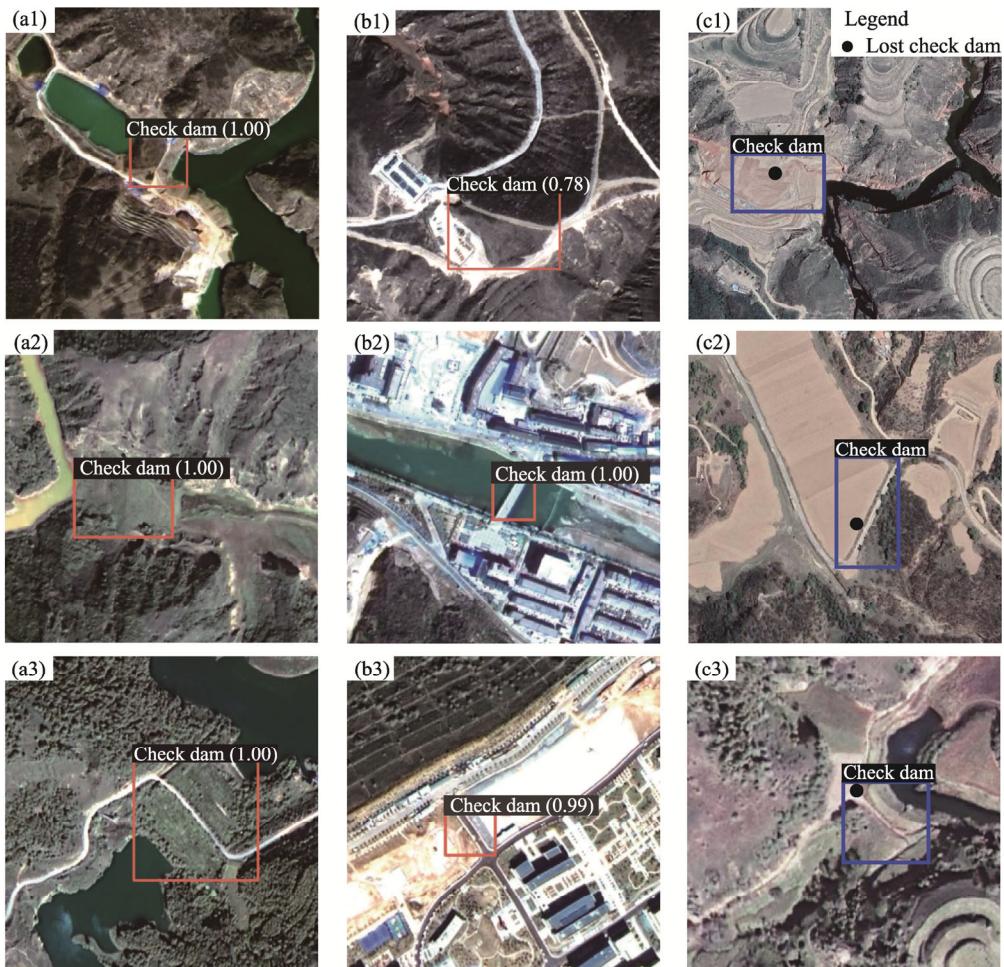
**Fig. 8** Spatial distribution of check dam density in the Yanhe River Basin (mapped using Kernel density with 8000 m bandwidth in ArcGIS 10.2)

## 4 Discussion

### 4.1 Performance of proposed framework in check dam detection

Traditional check dam monitoring mainly relies on manual surveys, which are time-consuming and labor-intensive, and the data can lack objectivity and accuracy (Chen and Zhang, 2004). Remote sensing technology has clear advantages over traditional methods. Tian et al. (2013) used remote sensing images in conjunction with a field survey to derive the spatial distribution of check dams in Huangfu Chuan River. However, they only extracted check dams or reservoirs with water bodies, ignoring check dams with other land covers such as cropland. Moreover, it is hard to obtain the number and distribution of check dams using their method. Compared with field surveys or other traditional image processing techniques, the frame proposed in this study can record the quantity and distribution of check dams in broad regions more objectively, timely, and effectively at a lower cost and avoid duplicating manual survey data. In addition, we evaluated five models to identify check dams and found that VFNet and YOLOX performed better than the other three models. VFNet performed best for detecting large check dams, while YOLOX performed best for detecting medium check dams. In practical applications, a relatively low probability threshold of 0.5 was used to detect as many check dams as possible, retaining predicted boxes with a confidence score greater than 0.5. However, this resulted in substantial overestimation, with many objects wrongly classified as check dams due to severe background interference and similar spectrum and texture characteristics in the GF-2 images between check dams and line-type buildings, such as bridges or roads (Fig. 9).

The deep learning models also failed to detect some check dams (Fig. 6c), particularly those built long ago and now filled with sediment. The GF-2 images only showed the top parts of these check dams, not their spectral or textural features. The shortage of training samples also resulted in some check dams not being recognized correctly. Therefore, strategies are needed to alleviate the abovementioned problems and refine the identification accuracy of check dams. Most studies have focused on improving the network structure to increase recognition accuracy, including R-CNN and YOLO networks (Sharma and Mir, 2020). Our study showed that Cascade R-CNN and YOLOX had greater detection accuracy than Faster R-CNN and YOLOv3 (Table 3), respectively. For remote sensing researchers, the post-processing method based on geospatial



**Fig. 9** Results of check dam detections based on deep learning models and lost check dams. (a1–a3), correctly detected check dams; (b1–b3), incorrect detections; (c1–c3), lost check dams. Detection results in (a1–a3) and (b1–b3) show the predicted bounding box of check dam and its corresponding confidence score.

analysis is an efficient attempt to improve object extraction results. The geospatial analysis can reflect the spatial distribution constraint relationship between the location of ground objects and certain geographic data, such as DEM and land cover. We selected channel buffer areas obtained from DEM and land cover as the restriction factor, eliminating misrecognized check dams, removing 50.0% false detection boxes, and improving the precision indicator by 9.0%. Similarly, Zeng et al. (2019) proposed an airport detection method using spatial analysis and deep learning. They first reduced the candidate airport regions to 0.56% of the total area of 75,691 km<sup>2</sup> based on spatial analysis of released remote sensing products, including global land cover (FROM-GLC10), ALOS Global Digital Surface Model (ALOS World 3D-30m), and open street map (OSM) datasets. Then, they used Faster R-CNN to determine the airport location and obtained a mean user's accuracy of 88.9%, ensuring that all aircraft could be detected. Zhang et al. (2022) used street view images to identify road noise barriers with deep learning classification models and geospatial analysis and acquired final road noise barrier identification results in Suzhou City of Jiangsu Province, China. However, the effectiveness of geospatial analysis relies heavily on the accuracy of geographical data. Applying DEM and land cover products in our research accumulated errors, impacting the precision of check dam identification. In order to minimize the effect of land cover on check dam recognition, we used ESA WorldCover at 10 m resolution for 2020, the same year as the applied GF-2 remote sensing images, which eliminated the effect of interannual change of land use.

Zanaga et al. (2021) showed that ESA WorldCover 10 m 2020 v100 captured landscapes at a higher level of detail than Environmental Systems Research Institute (ESRI) 2020 Landcover. Moreover, we focused on detecting large and medium check dams with body lengths greater than 50 m; as such, the resolution effect of the selected land cover product is insignificant. Using higher resolution and precise DEM and land cover products can improve the detection results of the method. Also, economic, cultural, geographical, and other factors between regions will affect landforms and check dam construction and distribution and should be considered when using the proposed geographic analysis method for improving the precision of check dam recognition in different regions.

#### 4.2 Limitations and implications

Deep learning is a promising method for remote sensing image analysis (Ghanbari et al., 2021). In this study, we integrated widely used object detectors with spatial analysis to explore the distribution of check dams at watershed scale. However, the proposed framework is subject to some limitations. It is hard to recognize small check dams due to the limited resolution of remote sensing images. In addition to the about 1100 large and medium check dams detected by our method, thousands of smaller check dams in the Yanhe River Basin were blurred and indistinguishable in the GF-2 images and thus not considered when preparing the sample dataset. Most small check dams were built by local farmers from 1950 to 1980 (Liu et al., 2018) to develop agricultural production (also referred to as 'production dams'). Therefore, we dismissed these small check dams due to their limited effectiveness in regulating runoff and controlling sediment. In addition, the trained models identified check dams in the Loess Plateau because we customized samples in this region. The construction of other check dams worldwide used various materials such as stones, earth, wood logs, and straw bales (Abbasi et al., 2019; Lucas-Borja et al., 2019; Robichaud et al., 2019). Collecting more samples of check dams made from various materials in different environments to enrich sample datasets is needed to extend the range of the proposed framework.

Combining multidisciplinary sciences such as remote sensing, deep learning, and geographic information system (GIS) is a lower-cost method for identifying check dams than field surveys and other traditional image processing methods. However, few studies have focused on check dam detection using deep learning and remote sensing methods. One study focused on extracting dam areas by integrating OBIA and the semantic segmentation approach (Li et al., 2021). The authors reported the feasibility of their method but only tested it in four small regions. Acquiring information on check dams from remote sensing imagery could be more comprehensive and accurate if we combine Li et al.'s method for check dam area extraction with our proposed method for dam body detection. Detailed information on check dams at the watershed scale, including their number, location, spatial distribution, and control area, can help analyze the effect of check dams on erosion reduction and plan suitable dam sites.

### 5 Conclusions

This study proposed a rapid and precise check dam identification method in broad areas from high-resolution remote sensing images using deep learning and geographic analysis. We compared five advanced deep learning object detectors, including Faster R-CNN, YOLOv3, Cascade R-CNN, YOLOX, and VFNet, with all performing well for detecting check dams. However, VFNet and YOLOX had more robust capabilities for check dam identification, with AP values greater than 69.0%, 96.0%, and 80.0% at IoU thresholds of 0.50:0.95, 0.50, and 0.75, respectively. We combined preferred deep learning detection models with geospatial analysis to identify check dams in the Yanhe River Basin; the precision and recall rates reached 87.6% and 86.5%, respectively. Moreover, the proposed method identified recently constructed dams not recorded in the survey data. Our method also identified the location and spatial distribution of check dams in the Yanhe River Basin, with regional differences in spatial distribution. The central and northeastern parts of the Yanhe River Basin are two agglomeration areas with a high density of check dams. We expect to use this method to detect check dams on a national scale.

## Acknowledgements

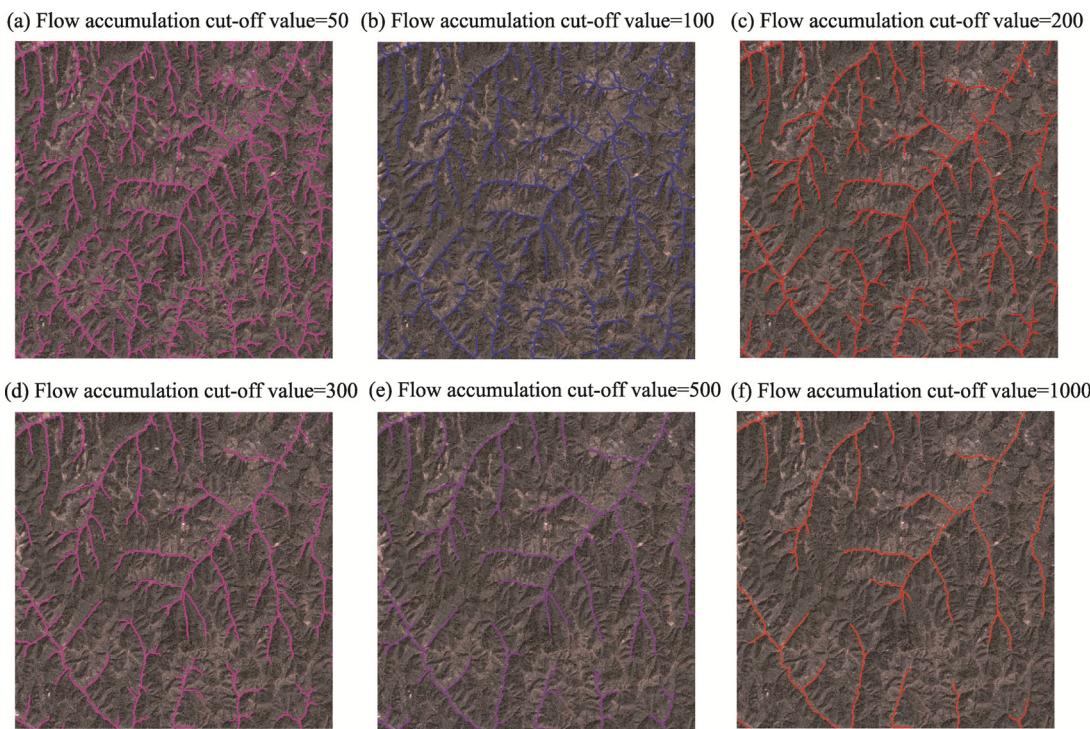
This research was supported by the National Natural Science Foundation of China (41977064) and the National Key R&D Program of China (2021YFD1900700). The authors express their gratitude to colleagues in their research group for helping complete the experiments.

## References

- Abbasi N A, Xu X Z, Lucas-Borja M E, et al. 2019. The use of check dams in watershed management projects: Examples from around the world. *Science of the Total Environment*, 676: 683–691.
- Alfonso-Torreño A, Gómez-Gutiérrez Á, Schnabel S, et al. 2019. sUAS, SfM-MVS photogrammetry and a topographic algorithm method to quantify the volume of sediments retained in check-dams. *Science of the Total Environment*, 678: 369–382.
- Apolo-Apolo O E, Martínez-Guanter J, Egea G, et al. 2020. Deep learning techniques for estimation of the yield and size of citrus fruits using a UAV. *European Journal of Agronomy*, 115: 126030, doi: 10.1016/j.eja.2020.126030.
- Bai L C, Wang N, Jiao J Y, et al. 2020. Soil erosion and sediment interception by check dams in a watershed for an extreme rainstorm on the Loess Plateau, China. *International Journal of Sediment Research*, 35(4): 408–416.
- Bai M, Mo X G, Liu S X, et al. 2019. Contributions of climate change and vegetation greening to evapotranspiration trend in a typical hilly-gully basin on the Loess Plateau, China. *Science of the Total Environment*, 657: 325–339.
- Cai Z W, Vasconcelos N. 2018. Cascade R-CNN: Delving into high quality object detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Salt Lake City, USA.
- Chen J Q, Zhang F. 2004. IKONOS secondary planet image data applied in a small watershed for soil and water conservation work. *Bulletin of Soil and Water Conservation*, 24(5): 64–66. (in Chinese)
- Chen K, Wang J Q, Pang J M, et al. 2019. MMDetection: Open MMLab Detection Toolbox and Benchmark. [2022-06-20]. <https://doi.org/10.48550/arXiv.1906.07155>.
- Chen Z, Zhang T, Ouyang C. 2018. End-to-End airplane detection using transfer learning in remote sensing images. *Remote Sensing*, 10(1): 139, doi: 10.3390/rs10010139.
- Ding P, Zhang Y, Deng W J, et al. 2018. A light and faster regional convolutional neural network for object detection in optical remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 141: 208–218.
- Fu K, Chen Z, Zhang Y, et al. 2019. Enhanced feature representation in detection for optical remote sensing images. *Remote Sensing*, 11(18): 2095, doi: 10.3390/rs11182095.
- Ge Z, Liu S T, Wang F, et al. 2021. YOLOX: Exceeding YOLO Series in 2021. [2022-06-20]. <https://doi.org/10.48550/arXiv.2107.08430>.
- Ghanbari H, Mahdianpari M, Homayouni S, et al. 2021. A meta-analysis of convolutional neural networks for remote sensing applications. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14: 3602–3613.
- Hou Q Q. 2013. Soil conservation practices information extraction from remote sensing images using object-oriented method. MSc Thesis. Yangling: Northwest A&F University. (in Chinese)
- Jin Z, Cui B L, Song Y, et al. 2012. How many check dams do we need to build on the Loess Plateau? *Environmental Science & Technology*, 46(16): 8527–8528.
- Kamilaris A, Prenafeta-Boldu F X. 2018. Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, 147: 70–90.
- Khelifi L, Mignotte M. 2020. Deep learning for change detection in remote sensing images: comprehensive review and meta-analysis. *IEEE Access*, 8: 126385–126400.
- Konstantinidis D, Argyriou V, Stathaki T, et al. 2020. A modular CNN-based building detector for remote sensing images. *Computer Networks*, 168: 107034, doi: 10.1016/j.comnet.2019.107034.
- Laben C A, Brower B V. 2000. Process for Enhancing the Spatial Solution of Multispectral Imagery Using Pan-Sharpening. [2022-06-20]. <https://www.freepatentsonline.com/6011875.html>.
- Li N, Cheng L, Huang L Y, et al. 2021. Framework for unknown airport detection in broad areas supported by deep learning and geographic analysis. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14: 6328–6338.
- Li S J, Xiong L Y, Hu G H, et al. 2021. Extracting check dam areas from high-resolution imagery based on the integration of object-based image analysis and deep learning. *Land Degradation & Development*, 32(7): 2303–2317.
- Lin T Y, Maire M, Belongie S, et al. 2014. Microsoft Coco: Common Objects in Context. [2022-06-21]. [https://doi.org/10.1007/978-3-319-10602-1\\_48](https://doi.org/10.1007/978-3-319-10602-1_48).
- Liu X Y, Gao Y F, Ma S B, et al. 2018. Sediment reduction of warping dams and its timeliness in the Loess Plateau. *Journal of Hydraulic Engineering*, 49(2): 145–155. (in Chinese)
- Liu Z. 2013. The national census for soil erosion and dynamic analysis in China. *International Soil and Water Conservation Research*, 1(2): 12–18.

- Lucas-Borja M E, Piton G, Nichols M, et al. 2019. The use of check dams for soil restoration at watershed level: a century of history and perspectives. *Science of the Total Environment*, 692: 37–38.
- Lucas-Borja M E, Piton G, Yu Y, et al. 2021. Check dams worldwide: Objectives, functions, effectiveness and undesired effects. *CATENA*, 204: 105390, doi: 10.1016/j.catena.2021.105390.
- Mahdianpari M, Salehi B, Rezaee M, et al. 2018. Very deep convolutional neural networks for complex land cover mapping using multispectral remote sensing imagery. *Remote Sensing*, 10(7): 1119, doi: 10.3390/rs10071119.
- Mi Z J, Mu X M, Zhao G J. 2015. Extraction of check dam based on multi-sources data in the Huangfuchuan watershed. *Arid Land Geography*, 38(1): 52–59. (in Chinese)
- Ministry of Water Resources of China. 2013. *Bulletin of First National Census for Water*. Beijing: China Water & Power Press. (in Chinese)
- Mur A, Travé-Massuyès L, Chanthery E, et al. 2022. A neural algorithm for the detection and correction of anomalies: Application to the landing of an airplane. *Sensors*, 22(6): 2334, doi: 10.3390/s22062334.
- Ngula Niipele J, Chen J P. 2019. The usefulness of ALOS-PALSAR DEM data for drainage extraction in semi-arid environments in the Iishana sub-basin. *Journal of Hydrology: Regional Studies*, 21: 57–67.
- Piton G, Recking A. 2017. Effects of check dams on bed-load transport and steep-slope stream morphodynamics. *Geomorphology*, 291: 94–105.
- Pourghasemi H R, Yousefi S, Sadhasivam N, et al. 2020. Assessing, mapping, and optimizing the locations of sediment control check dams construction. *Science of the Total Environment*, 739: 139954, doi: 10.1016/j.scitotenv.2020.139954.
- Rahmati O, Ghasemieh H, Samadi M, et al. 2020. TET: An automated tool for evaluating suitable check-dam sites based on sediment trapping efficiency. *Journal of Cleaner Production*, 266: 122051, doi: 10.1016/j.jclepro.2020.122051.
- Reda K, Kedzierski M. 2020. Detection, classification and boundary regularization of buildings in satellite imagery using faster edge region convolutional neural networks. *Remote Sensing*, 12(14): 2240, doi: 10.3390/rs12142240.
- Redmon J, Farhadi A. 2018. Yolov3: An Incremental Improvement. [2022-06-22]. <https://doi.org/10.48550/arXiv.1804.02767>.
- Ren S Q, He K M, Girshick R, et al. 2017. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6): 1137–1149.
- Robichaud P R, Storrar K A, Wagenbrenner J W. 2019. Effectiveness of straw bale check dams at reducing post-fire sediment yields from steep ephemeral channels. *Science of the Total Environment*, 676: 721–731.
- Sharma V, Mir R N. 2020. A comprehensive and systematic look up into deep learning based object detection techniques: A review. *Computer Science Review*, 38: 100301, doi: 10.1016/j.cosrev.2020.100301.
- Shi P, Zhang Y, Ren Z P, et al. 2019. Land-use changes and check dams reducing runoff and sediment yield on the Loess Plateau of China. *Science of the Total Environment*, 664: 984–994.
- Sun P C, Wu Y P. 2022. Dynamic Modeling Framework of Sediment Trapped by Check-dam Networks: A Case Study of a Typical Watershed on the Chinese Loess Plateau. [2022-06-23]. <https://doi.org/10.1016/j.eng.2021.12.015>.
- Tian P, Zhao G J, Mu X M, et al. 2013. Check dam identification using multisource data and their effects on streamflow and sediment load in a Chinese Loess Plateau catchment. *Journal of Applied Remote Sensing*, 7: 073697, doi: 10.1111/j.JRS.7.073697.
- Wang Y F, Fu B J, Chen L D, et al. 2011. Check dam in the Loess Plateau of China: Engineering for environmental services and food security. *Environmental Science & Technology*, 45(24): 10298–10299.
- Wang Z Y, Chen Z Y, Yu S, et al. 2021. Erosion-control mechanism of sediment check dams on the Loess Plateau. *International Journal of Sediment Research*, 36(5): 668–677.
- Wei Y H, He Z, Jiao J Y, et al. 2018. Variation in the sediment deposition behind check-dams under different soil erosion conditions on the Loess Plateau, China. *Earth Surface Processes and Landforms*, 43(9): 1899–1912.
- Wu W T, Liu H, Li L L, et al. 2021. Application of local fully Convolutional Neural Network combined with YOLO v5 algorithm in small target detection of remote sensing image. *PLoS ONE*, 16(10): e0259283, doi: 10.1371/journal.pone.0259283.
- Yazdi J, Sabbaghian Moghaddam M, Saghafian B. 2018. Optimal design of check dams in mountainous watersheds for flood mitigation. *Water Resources Management*, 32: 4793–4811.
- Zanaga D, Van De Kerchove R, De Keersmaecker W, et al. 2021. ESA WorldCover 10 m 2020 v100. [2022-06-23]. <https://viewer.esa-worldcover.org/worldcover>.
- Zeng F X, Cheng L, Li N, et al. 2019. A hierarchical airport detection method using spatial analysis and deep learning. *Remote Sensing*, 11(19): 2204, doi: 10.3390/rs11192204.
- Zhang H Y, Wang Y, Dayoub F, et al. 2021. VarifocalNet: An IoU-aware Dense Object Detector. [2022-06-22]. <https://doi.org/10.48550/arXiv.2008.13367>.
- Zhang K, Qian Z, Yang Y, et al. 2022. Using street view images to identify road noise barriers with ensemble classification model and geospatial analysis. *Sustainable Cities and Society*, 78: 103598, doi: 10.1016/j.scs.2021.103598.
- Zhao D B. 2007. The analysis and extraction of soil conservation practices information based on high-resolution remote sensing images. MSc Thesis. Yangling: Northwest A&F University. (in Chinese)

## Appendix



**Fig. S1** Gully networks extracted by different flow accumulation cut-off values. (a), flow accumulation cut-off value=50; (b), flow accumulation cut-off value=100, (c), flow accumulation cut-off value=200; (d), flow accumulation cut-off value=300; (e), flow accumulation cut-off value=500; (f), flow accumulation cut-off value=1000.